Are We Accepting the Right Students to Graduate Engineering Programs: Measuring the Success of Accepted Students via Data Envelopment Analysis

School of Engineering
University of Bridgeport
Bridgeport, CT 06604

ABSTRACT

In this paper, we present two consecutive DEA models to measure the relative efficiency of applicants to graduate programs in engineering and to compare these efficiencies with the success of these students in the program. The proposed performance criteria are determined depending on the current evaluation criteria in the School of Engineering at the University of Bridgeport. The steps and implementation of the proposed methodology are explained with the help of a numerical example for the Fall 2004 semester.

Keywords: Graduate Enrollment, Engineering, Decision Making, Engineering Education, Data Envelopment Analysis.

1. Introduction

Evaluating candidates for graduate degree programs has always been a concern of both academic and administrative personnel at Universities. The difficulty of this task has increased over time due to growing complexity and size of the pool of applicants as educational programs extend to the global arena. Many Universities are facing a significant increase in the number of international student applications to graduate degree programs.

With this being the motivation, this study aims at determining the key criteria for applicants to the graduate programs at the University of Bridgeport, School of Engineering. In this regard, a two-step approach is developed. In the first step, an output oriented Data Envelopment Analysis (DEA) has been utilized to evaluate and rank the accepted applicants depending on various criteria; for example, GRE and TOEFL scores, GPA, number of below-B grades in the Bachelor of Science transcripts, and other parameters. Following this, an additional ranking algorithm is implemented and run to determine the degree of success among the same set of accepted students, following their progress in the program till they graduate.

The results of the two ranking algorithms are then compared to validate the appropriateness of the selection criteria. A case study is included to demonstrate the steps and applicability of the proposed DEA approach.

Data envelopment analysis (DEA) is a widely applied linear programming-based technique first developed by Charnes *et al.*[1] in 1978 to evaluate the efficiency of a set of decision-making units. Since then DEA has mostly been used for benchmarking and for performance evaluation purposes.

The paper is organized as follows: A brief list of previous studies is summarized in the following section. Section 3 provides a summary of the Data Envelopment Analysis approach. The problem description and a case study are the focus of Section 4. Conclusions and thoughts for future research are then provided in Section 5.

2. Literature review

Data Envelopment Analysis (DEA) is a non-parametric approach that compares similar entities, i.e., decision making units (DMUs), against the "best virtual decision making unit". The approach has the ability to accommodate multiple inputs and multiple outputs, allowing these variables to be included in the model with different units of measurement. Due to these advantages and ease in its use, the approach has been employed extensively in various areas, such as health care, education, banking, manufacturing, and management.

One of the most relevant studies is published by Johnson and Zhu [2]. In their work, the authors employed DEA to select the most promising candidates to fill an open faculty position. In this regard, the authors proposed a DEA aided recruiting process that (1) determines the performance levels of the ''best'' candidates relative to other applicants; (2) evaluates the degree of excellence of ''best'' candidates' performance; (3) forms consistent tradeoff information on multiple recruiting criteria among search committee members, and, then, (4) clusters the applicants.

DEA also found a large variety of applications in the environmental arena. To this extend, Sarkis [3] proposed a two-stage methodology to integrate managerial preferences and environmentally conscious manufacturing (ECM) programs. Consequently, Sarkis and Cordeiro [4] investigated the relationship between environmental and financial performance at the firm level.

Furthermore, Talluri *et al*. [5] applied DEA and Goal Programming methods to a Value Chain Network (VCN) considering the cross efficiency evaluations of Decision Making Units (DMUs).

Methods other than DEA have also been utilized to study the efficiency of application and admission processes. Moore [6] built an operational two-stage expert system to examine the admission decision process for applicants to an MBA program, and predict the degree completion potential for those actually admitted. A similar study is also published by Nilsson [7] to investigate any differences in the predictive relationships between the scores of the Graduate Record Examination (GRE), the graduate grade point average, the scores of the Graduate Management Admission Test (GMAT), and the graduate grade point average. Furthermore, Landrim *et al*. [8] constructed a value tree diagram for fifty-five graduate institutions offering the Ph.D. degree in psychology. The authors made use of this diagram to indicate the relative weight of admission factors used in the decision making process.

This study is a follow up on Kongar and Sobh's[9] previously published work where the authors proposed a DEA approach to measure the relative efficiency of applicants to the graduate programs in engineering. The proposed performance criteria in the study were determined depending on the current evaluation criteria in the School of Engineering at the University of Bridgeport.

3. Introduction to the data envelopment analysis approach

Data Envelopment Analysis (DEA) is a non-parametric approach that compares similar entities, i.e., decision making units (DMUs), against the "best virtual decision making unit". Usually modeled as a linear programming (LP) model, the method provides relative efficiency score for each decision making unit under consideration.

The most appealing advantage of DEA is, unlike parametric approaches such as regression analysis (RA), DEA optimizes on each individual observation not requiring a single function that suits best to all observations (Charnes *et al.*[10]). Comparison of DEA and RA has been well studied in the literature. Majority of the published work accept that DEA is more advantageous in comparing decision making units even though there are some studies emphasizing the advantages of both (*i.e.*, see Thanassoulis[11]).

One of the above mentioned comparative studies is published by Banker *et al.* [12], comparing estimates of technical efficiencies of individual hospitals obtained from the econometric modeling of the translog cost function, and the application of DEA. In their study, the authors reported that DEA estimates were highly related to the capacity utilization, whereas translog estimates failed to provide such relationship.

In addition, Bowlin *et al.* [13] compared DEA and RA via 15 hypothetical hospitals and concluded that DEA outperformed RA with its ability to identify the sources of inefficiencies by underlining the resources that are used in excess in inefficient hospitals. Furthermore, the authors stated that DEA performed superior in estimating and returning scale characterizations. In addition, Sarkis [14] compared DEA and conventional multiple criteria decision making (MCDM) tools in terms of efficiency and concluded that DEA appeared to perform well as a discrete alternative MCDM tool.

DEA algorithms can be classified into two categories, *input-* and *output-oriented* DEA models, according to the "orientation" of the model. *Input-oriented* DEA models concentrate on reducing the amount of input by keeping the output constant. *Output-oriented* DEA models on the other hand, focus on maximizing the amount of output with the identical amount of input. In DEA modeling, inputs are considered as the items that are subject to minimization (*i.e.*, less is better), whereas, outputs are the items that are subject to maximization (*i.e.*, more is better).

Further classification of DEA models can be given depending on the "optimality scale" criterion. Here, DEA models can work under the assumption of Constant Returns to Scale (CRS), or non-constant returns to scale, *i.e.*, Increasing Returns to Scale (IRS), "Decreasing Returns to Scale (DRS)", and "Variable Returns to Scale (VRS)"; implying that not all DMUs are functioning at a optimality scale. Here, CRS assumes changes in output values subsequent to a proportional change in the input values. VRS was initially introduced by Banker et al. 15 as an extension of the CRS DEA model. In this paper, we employ an output oriented CRS DEA model. Further explanation regarding the CRS model follows.

As also mentioned above, a basic DEA model allows the introduction of multiple inputs and multiple outputs and obtains an "efficiency score" of each DMU with the conventional output/input ratio analysis. Defining basic efficiency as the *ratio of weighted sum of outputs to the weighted sum of inputs*, the relative efficiency score of a test DMU p can be obtained by solving the following DEA ratio model (CCR) proposed by Charnes, *et al.*[1]:

$$
\begin{aligned}
& max \quad \frac{\sum_{k=1}^{s} v_k y_{kp}}{\sum_{j=1}^{m} u_j x_{jp}} \\
& s.\ t. \quad \frac{\sum_{k=1}^{s} v_k y_{ki}}{\sum_{j=1}^{m} u_j x_{ji}} \leq 1 && \forall \text{ DMUs } i \\
& \qquad v_k, u_j \geq 0 && \forall\ k, j.
\end{aligned} \qquad (1)
$$

Where,
$k = 1$ to s,
$j = 1$ to m,
$i = 1$ to n,
y_{ki} = amount of output k produced by DMU i,
x_{ji} = amount of input j produced by DMU i,
v_k = weight assigned to output k,
u_j = weight assigned to input j.

Equation (1) can be easily converted into a linear program as in Equation (2). We refer the reader to the study by Charnes *et al.*[10] for further explanation of the model.

$$
\begin{aligned}
& max \quad \sum_{k=1}^{s} v_k y_{kp} \\
& s.\ t. \quad \sum_{j=1}^{m} u_j x_{jp} = 1 \\
& \qquad \sum_{k=1}^{s} v_k y_{ki} - \sum_{j=1}^{m} u_j x_{ji} \leq 0 && \forall \text{ DMUs } i \\
& \qquad v_k, u_j \geq 0 && \forall\ k, j,
\end{aligned} \qquad (2)
$$

where, the $\sum_{j=1}^{m} u_j x_{jp} = 1$ constraint sets an upper bound of 1 for the relative efficiency score.

In the CCR model provided in Equation (2), evaluating the efficiency of n DMUs correspond to a set of n LP problems. Using duality, the dual of the CRS model can be represented as in Eq. (3):

$$
\begin{aligned}
& min \quad \theta \\
& s.t. \\
& \qquad \sum_{i=1}^{n} \lambda_i x_{ji} - \theta x_{jp} \le 0 \qquad && \forall \text{ Inputs } j \\
& \qquad \sum_{i=1}^{n} \lambda_i y_{ki} - y_{kp} \ge 0 \qquad && \forall \text{ Outputs } k \qquad (3) \\
& \qquad \lambda_i \ge 0 \qquad && \forall \text{ DMUs } i.
\end{aligned}
$$

Equation 3 corresponds to the dual of the basic input-oriented CCR model assuming constant returns to scale for all the inputs and outputs. Using Talluri's [16] notation, the dual of a basic output-oriented CRS model can be written as follows:

$$
\begin{aligned}
& max \quad \phi \\
& s.t. \\
& \qquad x_{jp} - \sum_{i} \lambda_i x_{ji} \ge 0 \qquad && \forall \text{ Inputs } j \\
& \qquad -\phi y_{kp} + \sum_{i} \lambda_i y_{ki} \ge 0 \qquad && \forall \text{ Outputs } k \qquad (4) \\
& \qquad \lambda_i \ge 0 \qquad && \forall \text{ DMUs } i.
\end{aligned}
$$

In the case where the assumption that not all DMUs are functioning at an optimality scale, Equation 4 could be converted into a VRS model by including the constraint $\sum_i \lambda_i \ge 0$ to the set of technological constraints.

The result of the model, Φ is the relative efficiency score of each DMU. Inverse of the variable Φ ($1/\Phi$) provides the technical efficiency value (TE) for each DMU. Here, given the technical efficiency value is equal to one ($TE = 1$), DMU p is considered "efficient" for its selected weights. In this case, DMU p lies on the optimal frontier and is not dominated by any other DMU. Using similar reasoning, if the technical efficiency value is less than one ($TE < 1$), then it can be claimed that DMU p is not on the optimal frontier and there exists at least one efficient DMU in the population.

The following demonstrates the application of the CRS DEA model to the evaluation process of the applicants for graduate engineering programs.

4. Applying data envelopment analysis to the application review process

The proposed DEA model in this study aims at (i) accepting students, (ii) comparing the accepted students with the DEA model results, and, (iii) preparing a base to observe the students' future success to evaluate the performance criteria fed into the model.

To achieve these objectives, the data for all 37 M.S. candidates (n = 37) for the Masters of Science (M.S.) in Computer Science program in the School of Engineering for Fall 2004 semester is collected.

After reading in the relevant data, a DEA model is employed to evaluate the relative efficiency of each candidate using six performance criteria, viz., the Bachelors of Science (B.S.) GPA (BS GPA), TOEFL and GRE Quantitative (GRE-Q) scores, number of years of work experience, number of undergraduate semesters till B.S. degree completion, and the number of below-B grades in math-related and technical courses in the B.S. degree transcript.

4.1 DEA model for the evaluation process

Following the retrieval of the complete application materials, related data is entered into the applications database. The office of admissions then sends each applicant a confirmation e-mail with an assigned University of Bridgeport (UB) identification number confirming that the application has been received.

Subsequently, the applications are filtered by the office of admissions depending on basic application criteria, filtering out unqualified applicants. These applicants are then notified regarding the result of their applications. Remaining applications which meet the basic requirements are then sent to the relevant Faculty for decision making.

The information provided by this study enables users to identify the best candidates for the graduate engineering program. In the following sections, we illustrate how the evaluation process can be enhanced using the DEA approach introduced earlier.

4.2 DEA model I to evaluate the efficiency of candidates for graduate study

In our model, the applications to the graduate program correspond to decision-making units in DEA, while application data correspond to criteria in DEA, dependent on the definition of the indicators (inputs or outputs in the DEA model).

In total, the model embodies 37 decision-making units and six criteria. These criteria include two inputs and four outputs. Input criteria consist of e_1, and e_2, whereas output criteria include, e_3, e_4, e_5, and, e_6, where,

e_1 = number of below-B grades in math-related/technical courses in the BS transcript of the applicant,
e_2 = number of semesters that the applicant spent to complete the BS degree,
e_3 = BS GPA of the applicant,
e_4 = TOEFL score of the applicant,
e_5 = GRE-Q score of the applicant,
e_6 = number of years of work experience of the applicant.

The first input introduced to the model is the number of below-B grades in math-related/technical courses in the B.S. transcript (e_1). Following the notation of the first DEA model, the first input formulation for each DMU i (x_{1i}) can be written as follows:

$$x_{1i} = e_{1i} \qquad \forall \text{ DMUs } i. \qquad (5)$$

The second input introduced to the model is the number of semesters spent to complete the B.S. degree, (e_2). Hence, the second input formulation for each DMU i (x_{2i}) can be written as follows:

$$x_{2i} = e_{2i} \qquad \forall \text{ DMUs } i. \qquad (6)$$

The output variables in the proposed DEA model are selected as, the B.S. GPA of the applicant (e_3), the TOEFL score of the applicant (e_4), the GRE-Q score of the applicant (e_5), and the number of years of previous work experience (e_6) of the applicant.

Therefore, with similar reasoning, equations (7), (8), (9), and (10) can be expressed mathematically as follows:

$$y_{1i} = e_{3i} \qquad \forall \text{ DMUs } i. \qquad (7)$$

$$y_{2i} = e_{4i} \qquad \forall \text{ DMUs } i. \qquad (8)$$

$$y_{3i} = e_{5i} \qquad \forall \text{ DMUs } i. \qquad (9)$$

$$y_{4i} = e_{6i} \qquad \forall \text{ DMUs } i. \qquad (10)$$

This completes the formulation of the DEA model. Selected application data for a total of 37 candidates are provided in Table 1.

Table 1. Initial data for the DEA model I

DMU #	e_1	e_2	e_3	e_4	e_5	e_6	DMU #	e_1	e_2	e_3	e_4	e_5	e_6
1	8	8	3.22	477	640	0	20	17	8	3.11	560	610	0
2	11	8	3.2	507	770	0	21	12	8	3.32	610	730	0
3	0	8	2.37	574	693	0	22	6	6	3.68	574	693	2
4	5	6	3.14	490	750	0	23	0	6	3.4	574	693	5
5	0	8	3.98	553	800	0	24	12	8	3.24	577	730	0
6	18	8	2.92	677	790	1	25	9	8	3.04	583	580	0
7	20	10	2.97	633	780	0	26	0	8	2.97	560	760	0
8	8	8	3.1	563	660	2	27	14	8	3.03	550	730	0
9	2	8	3.56	593	800	0	28	7	8	3.34	560	640	0
10	23	8	2.98	523	660	2	29	9	8	3.34	550	620	0
11	15	8	3.24	563	700	0	30	11	8	3.07	647	630	0
12	0	6	3.77	597	600	0	31	7	8	3.52	563	670	0
13	6	8	3.41	593	660	0	32	1	6	3.38	653	760	7
14	1	8	3.85	600	770	0	33	3	8	3.67	560	610	0
15	11	8	3.33	550	570	0	34	2	6	3.5	574	693	8
16	1	8	3.68	480	693	2.5	35	0	8	3.44	587	770	0
17	0	6	4	603	660	0	36	10	8	3	567	540	0
18	1	8	3.92	643	800	0	37	18	8	2.57	547	670	0
19	9	8	3.37	627	710	0	*Ave.*	*7.5*	*7.7*	*3.3*	*574.1*	*692.8*	*0.8*

Using this data set, the output-oriented DEA model is run for each applicant in the sample using DEA-Solver-PRO 5.0. DEA-Solver-PRO is a DEA software designed on the basis of the textbook by Cooper et al.[17] to solve and analyze DEA models. The results of the model are presented in Table 2 in descending order of *TE* I values.

Table 2. Relative efficiency score (*TE* I) and rank of each candidate

Rank	DMU#	*TE* I	Rank	DMU#	*TE* I
1	34	1.000	20	21	0.727
1	32	1.000	21	27	0.720
1	23	1.000	21	24	0.720
1	17	1.000	23	31	0.711
5	12	0.990	24	13	0.703
6	4	0.987	25	11	0.703
7	22	0.986	26	33	0.694
8	5	0.868	27	28	0.677
9	35	0.833	28	25	0.671
10	18	0.823	29	8	0.667
11	26	0.823	30	1	0.666
12	14	0.799	31	29	0.666
13	9	0.790	32	15	0.663
14	6	0.780	33	37	0.661
15	2	0.760	34	20	0.657
16	3	0.750	35	36	0.655
17	30	0.743	36	10	0.655
18	16	0.739	37	7	0.616
19	19	0.728	Average		0.774

According to the DEA results depicted in Table 2, Candidates 34, 32, 23, and 17 are efficient in terms of their pre-application academic performances with technical efficiency (*TE* I) values equal to 1. All other applicants have a potential to increase the relative efficiency of academic performances by 1 minus the *T*E value. For instance, the efficiency of candidate 20 could be increased by 34.3%. The two lowest technical efficiency values are calculated for Candidates 36, 10 and 7 with 65.5%, 65.5% and 61.6%, respectively.

These low values are most probably driven by the high numbers of below-B grades in math-related/technical courses in the BS transcript (10, 23, and 20, respectively) and the low GPAs of the applicants (3, 2.98, and 2.97, respectively).

The average efficiency for the sample is 77.4%. Figure 1 represents the average efficiency and the *TE* I values for the 37 candidates in the population. As illustrated by Figure 1, 23 candidates fall below the average efficiency value (app. 62% of the candidates).

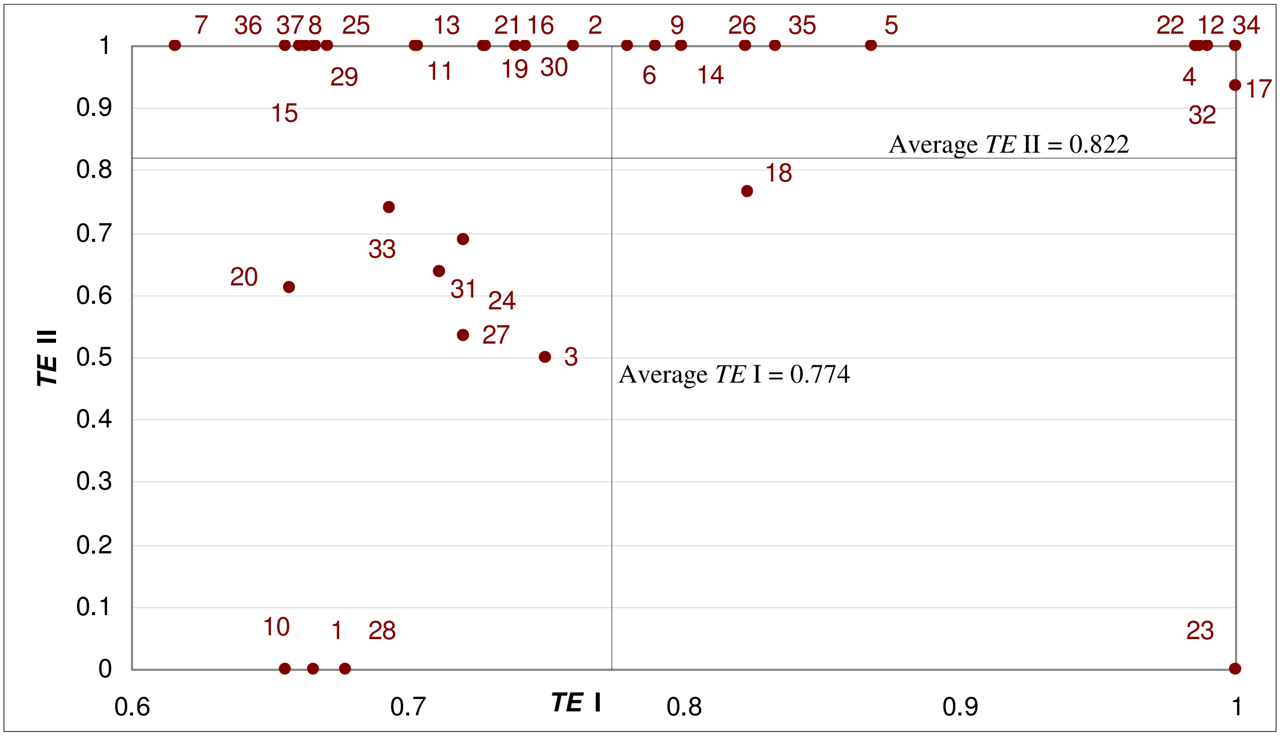

Figure 1. Performance efficiencies of 37 candidates according to the DEA I and DEA II model results.

As we analyze the results further, we can easily observe that all of the efficient candidates have completed their B.S. degrees in an identical number of semesters (6). In addition, the efficient candidates are characterized by either significantly high GPAs, GRE-Q scores, years of work experience, significantly low numbers of below-B grades in math-related/technical courses, or a combination of these criteria.

With this in mind, depending on the importance of each criterion, the input data can be normalized and weighed according to the decision maker preferences, so that the more important criterion would provide competitive advantage to the candidate.

In the following, a subsequent DEA model (DEA model II) is proposed to measure the relative efficiency of the future success of M.S. candidates.

4.3 DEA model II to evaluate the efficiency of candidates for graduate study

In this section, an output-oriented DEA model (DEA model II) is constructed to seek a relationship between the relative efficiency measures of the graduate students and their success in the graduate program.

In total, 37 decision-making units and four criteria are introduced. These criteria include three inputs and one output. The input criteria include t_1, whereas the output criteria include, t_2, t_3, and, t_4, where,

t_1 = number of below-C grades in the M.S. transcript of the M.S. candidate,
t_2 = GPA of the M.S. candidate,
t_3 = application status for the Curricular Practical Training (CPT) or Optional Practical Training (OPT) programs for the M.S. candidate; indicating whether they have applied to an industry internship during the program or a full-time position immediately following graduation,
t_4 = graduation status of the M.S. candidate.

The first input introduced to the DEA model II is the number of below-C grades in the M.S. transcript (t_1). Following the notation of the first DEA model, the first input formulation for each DMU i (x_{1i}) can be written as follows:

$$x_{1i} = t_{1i} \qquad \forall \text{ DMUs } i. \qquad (11)$$

The output variables in the proposed DEA model are selected as, the GPA of the M.S. candidate (t_2), the application status for CPT or OPT of the M.S. candidate (t_3), and the graduation status for of the M.S. candidate.

Therefore, with similar reasoning, equations (12), (13), and (14) can be expressed mathematically as follows:

$$y_{1i} = t_{2i} \qquad \forall \text{ DMUs } i. \qquad (12)$$
$$y_{2i} = t_{3i} \qquad \forall \text{ DMUs } i. \qquad (13)$$
$$y_{3i} = t_{4i} \qquad \forall \text{ DMUs } i. \qquad (14)$$

Here, for the application status for CPT or OPT of the M.S. candidate (y_{2i}); a positive integer value, ‘2’, is assigned if the M.S. candidate has applied for either CPT or OPT, where as the remaining variables are assigned the value of “1”.

Similar logic has been applied to the graduation status of the M.S. candidate (y_{3i}) and a positive integer value, “2”, is assigned if the M.S. candidate has graduated from the graduate degree program, where as “1” is assigned if the student has transferred out or if s/he is currently enrolled, but did not graduate yet.

The application data for a total of 37 candidates are depicted in Table 3.

Table 3. Initial data for the DEA model II

DMU #	t_1^*	t_2	t_3	t_4	DMU #	t_1	t_2	t_3	t_4
1	1	3.12	2	2	20	0	2.34	1	1
2	0	3.21	2	2	21	0	3.42	2	2
3	0	0.00	1	1	22	0	3.38	2	2
4	0	3.03	2	1	23	3	2.07	1	1
5	0	4.00	1	1	24	0	2.67	1	1
6	0	3.58	2	2	25	0	3.58	2	2
7	0	3.49	2	2	26	0	3.24	2	2
8	0	3.56	2	1	27	0	2.00	1	1
9	0	3.46	1	2	28	2	0.00	1	1
10	2	2.40	1	1	29	0	3.14	2	2
11	0	3.18	2	2	30	0	3.43	1	2
12	0	3.27	2	2	31	0	2.45	1	1
13	0	3.30	2	2	32	0	3.72	1	1
14	0	3.45	2	2	33	0	2.89	1	1
15	0	3.11	2	2	34	0	3.37	2	2
16	0	3.21	2	2	35	0	3.70	2	2
17	0	3.58	2	2	36	0	3.15	1	2
18	0	3.00	1	1	37	0	3.58	2	2
19	0	3.43	2	2	*Ave.*	0.2	3.01	1.6	1.6

*All zero values are changed to a significantly low positive value of 10^{-5} to avoid division by zero.

The results of the model are presented in Table 4 in descending order of *TE* II values.

Table 4. Relative efficiency score (*TE* II) and rank of each candidate

Rank	DMU#	*TE* II	Rank	DMU#	*TE* II
1	37	1	1	19	1
1	36	1	1	29	1
1	2	1	1	21	1
1	35	1	1	22	1
1	4	1	1	26	1
1	5	1	1	25	1
1	6	1	26	32	0.936
1	7	1	27	18	0.768
1	8	1	28	33	0.742
1	9	1	29	24	0.690
1	34	1	30	31	0.639
1	11	1	31	20	0.613
1	12	1	32	27	0.535
1	13	1	33	3	0.500
1	14	1	34	1	$1.0\text{x}10^{-5}$
1	15	1	35	10	$3.1\ \text{x}10^{-6}$
1	16	1	36	28	$2.5\ \text{x}10^{-6}$
1	17	1	37	23	$1.8\ \text{x}10^{-6}$
1	30	1	Average		0.822

According to the DEA results depicted in Table 4, Twenty five candidates are efficient in terms of their post-application academic performances, with technical efficiency (*TE* II) values equal to 1. All other applicants have a potential to increase the relative efficiency of academic performances by 1 minus the *T*E II value.

These low values are most probably driven by the lack of OPT or CPT applications and failure to graduate.

The average efficiency for the sample is 82.2%. Figure 1 represents the average efficiency and the *TE* II values for the 37 candidates in the population. As illustrated by Figure 1, 11 candidates fall below the average efficiency value.

Furthermore, it is difficult to establish a straight-forward or an obvious emerging pattern between the pre- and post-application relative efficiencies. However, we can observe that the proposed two-step DEA approach is more successful in determining the success of the applicants and is opposed to failures.

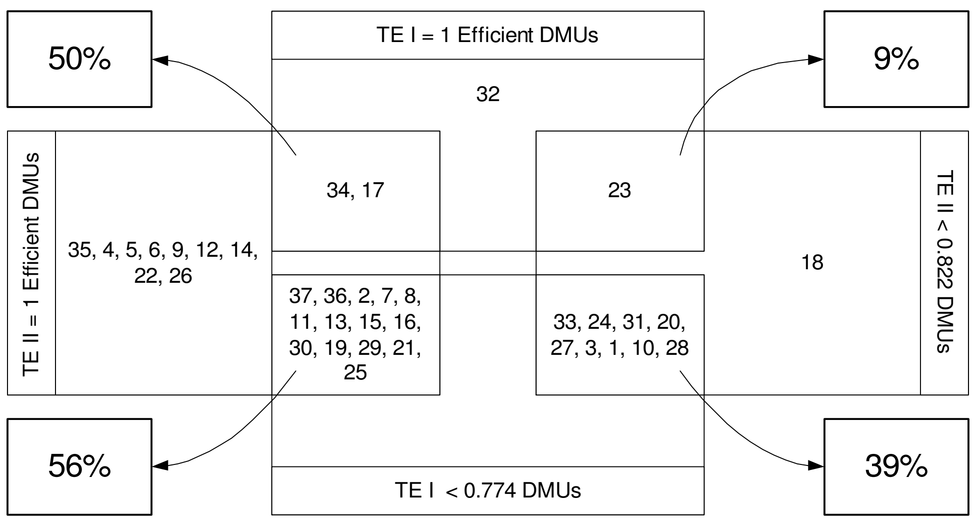

Figure 2. Results summary of the DEA models I and II.

Figure 2 illustrates detailed results of the proposed DEA model. As can be seen from Figure 2, in the proposed two step DEA model 50% of the DMUs fall into the intersection of efficient DMUs according to their TE I and TE II values. In addition, there is a 39% commonality for the DMUs that are below the average in both DEA models. Hence, we can conclude that the proposed DEA model is able to detect the efficient DMU more successfully compared to the ones that are below the average.

5. Conclusions and future research

In this study, implementations of two output-oriented DEA models are considered and applied to a sample of 37 accepted M.S. candidates to the Computer Science graduate degree program at the University of Bridgeport to determine the relative efficiency score of applicants based on their credentials. The model provides a basis to conduct a fast and reliable automated application evaluation process.

This is a follow-up study on past research conducted at the University of Bridgeport where we looked at the overall applicants to the graduate school and compared the relative efficiencies of candidates to manually selected one[9]. According to that study, we concluded that there was significant difference between the manually accepted candidates and the candidates ranked according to the DEA model results. This was most likely caused by (i) the inconsistency of the manual evaluation process and/or (ii) the presence of factors that are not included in the model; for example: the ranking of the university providing the B.S. degree, the B.S. major, the strength of the recommendation letters, etc.

In this study, we looked at the accepted candidates and analyzed their future performance to seek a correlation between the students' performance in the graduate program after admission and to compare the existing evaluation results, towards the eventual implementation of an automated graduate application admission system.

Both DEA steps proposed in the paper utilize the data for students who are both accepted and enrolled in the graduate engineering program. However, a considerable portion of accepted students, approximately 70-75%, do not enroll in the degree program even though they are accepted. This is due to visa acquisition problems and/or personal preference in attending a different University. Furthermore, data for rejected students are either not available or not reliable due to the recording and privacy laws limitations. Hence, recording the applications to the school and tracking each application so that the data set will include every student who applied to the program would surely provide much more reliable results.

In addition to the criteria utilized in the second DEA model, the duration of study, number of total credits and courses completed, and the numbers of grades less then C were also available; but these data were omitted, as they were considered to be not correlated with the graduate GPA and graduation status; which we considered, for the purpose of this study, to be the main indicators of success within the graduate course of study.

Furthermore, applications to OPT or CPT does not necessarily in all cases imply that the M.S. candidate has been employed by an organization. It only shows the intention of the M.S. candidate to seek employment in the U.S. after graduation. The employment data cannot be obtained in a reliable manner since keeping track of the employment status of graduate students is often difficult to accomplish in a timely manner.

In summary, the quality of the data highly affects the outcome of the proposed models. In the future, we are planning to collect the data solely for this purpose and track students from the application stage and follow their progress till they graduate. We plan to perform more correlation studies between the admission and graduate performance models and vary/change the number of the parameters for both models in order to fine-tune our system; towards the eventual goal of successfully implementing a fully-automated graduate admission system.

Acknowledgements

The authors would like to acknowledge the significant contributions, effort and support provided by Bryan Gross and Isabella Varga, of the Office of Admissions, and Yumin Wang, Director, International Affairs Office, at the University of Bridgeport.

REFERENCES

[1] A. Charnes, W. Cooper, and E. Rhodes, *Measuring the Efficiency of Decision-Making Units.* European Journal of Operational Research, 1978. **2**(6): p. 429–444.

[2] S. A. Johnson and J. Zhu, *Identifying "Best" Applicants in Recruiting Using Data Envelopment Analysis.* Socio-Economic Planning Sciences, 2003. **37**(2): p. 125-139.

[3] J. Sarkis, *A Methodological Framework for Evaluating Environmentally Conscious Manufacturing Programs.* Computers & Industrial Engineering, 1999. **36**(4): p. 793-810.

[4] J. Sarkis and J. J. Cordeiro, *An Empirical Evaluation of Environmental Efficiencies and Firm Performance: Pollution Prevention Versus End-of-Pipe Practice.* European Journal of Operational Research, 2001. **135**(1): p. 102-113.

[5] S. Talluri, R. C. Baker, and J. Sarkis, *Framework for Designing Efficient Value Chain Networks.* International Journal of Production Economics, 1999. **62**(1-2): p. 133-144.

[6] J. S. Moore, *An expert system approach to graduate school admission decisions and academic performance prediction.* Omega, 1998. **26**(5): p. 659-670.

[7] J. E. Nilsson, *The GRE and the GMAT: A Comparison of Their Correlations to GGPA.* Educational and Psychological Measurement, 1995. **55**: p. 637-40.

[8] R. E. Landrim, E. B. Jeglum, and J. R. Cashin, *The Decision-Making Process of Graduate Admissions Committees in Psychology.* Journal of Social Behavior and Personality, 1994. **9**(2): p. 239-248.

[9] E. Kongar and T. Sobh. *Evaluating the Efficiency of Candidates for Graduate Study via Data Envelopment Analysis*. in *2007 ASEE Annual Conference & Exposition*. 2007. Honolulu, Hawaii.

[10] A. Charnes, W. W. Cooper, A. Y. Lewin, and L. M. Seiford, *Data Envelopment Analysis: Theory, Methodology, and Applications*. 1994, Boston: Kluwer.

[11] E. Thanassoulis, *A Comparison of Regression Analysis and Data Envelopment Analysis as Alternative Methods for Performance Assessments.* The Journal of the Operational Research Society, 1993. **44**(11): p. 1129-1144.

[12] R. D. Banker, R. F. Conrad, and R. P. Strauss, *A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production.* Management Science, 1986. **32**(1): p. 30-44.

[13] W. F. Bowlin, A. Charnes, W. W. Cooper, and H. D. Sherman, *Data Envelopment Analysis and Regression Approaches to Efficiency Estimation and Evaluation.* Annals of Operations Research, 1984. **V2**(1): p. 113-138.

[14] J. Sarkis. *Ecoefficiency: How Data envelopment Analysis Can Be Used by Managers and Researchers*. in *Environmentally Conscious Manufacturing*. 2000. Boston, MA: Society of Photo-Optical Instrumentation Engineers.

[15] R. D. Banker, A. Charnes, and W. W. Cooper, *Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis.* Management Science, 1984. **30**(9): p. 1078-1092.

[16] S. Talluri, *Data Envelopment Analysis: Models and Extensions. Decision Line*, 2000. **31**(3): p. 8-11.

[17] W. W. Cooper, L. M. Seiford, and K. Tone, *Data Envelopment Analysis-A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. 2000: Springer.